Project introduction

Ideas and Wrap-up

Modalities

- Motivation: Get an idea how to apply the presented methods in-situ **or** how an integrative structural biology project is executed
- Time invested: Only the next 3.5 session per 2 hrs + 2 hrs of prep work each
- Mini-Symposium on 17.12. with 10 min presentation per group (starting at 8:00)
- Deliverables: 2 page report + presentation
- Possibility to work on izar (there is if free compute for this course)
- By today, decide on a project that I can help you individually

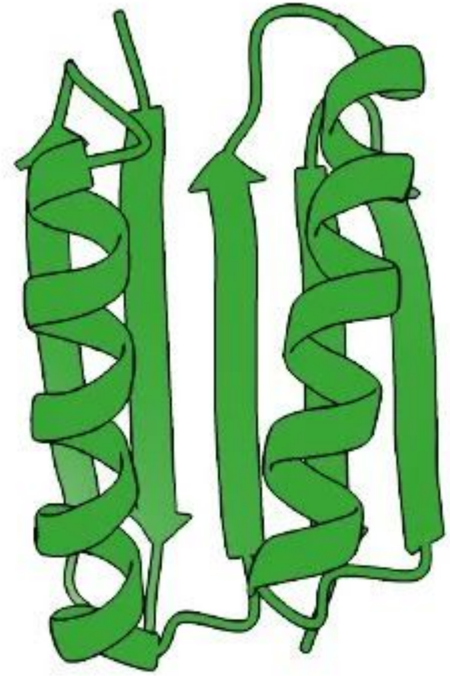
A small wrap-up

Mostly Protein Design related

What is Protein Design - then

- backbone design - heuristics based
- sequence assignment - painstakingly slow
- validation - physics based
- 100s - 1,000s of CPU hours
- months for a working design

Rosetta Commons

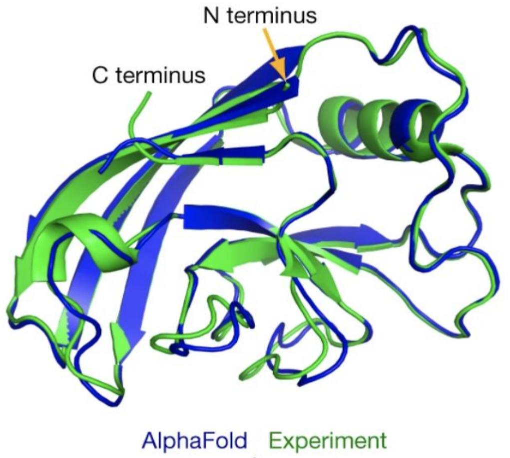


©Terezia Kovalova/The Royal Swedish Academy of Sciences

Advances in deep learning 🤝 protein design 1/2

The protein folding revolution: AlphaFold2

- General solution to protein structure prediction
- Predicts accurate protein structures in minutes
  - As opposed to months - years for experimental determination



r.m.s.d.$_{95}$ = 0.8 Å; TM-score = 0.93

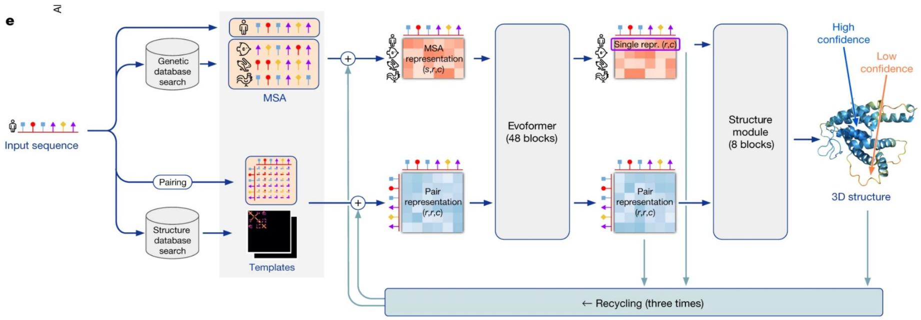
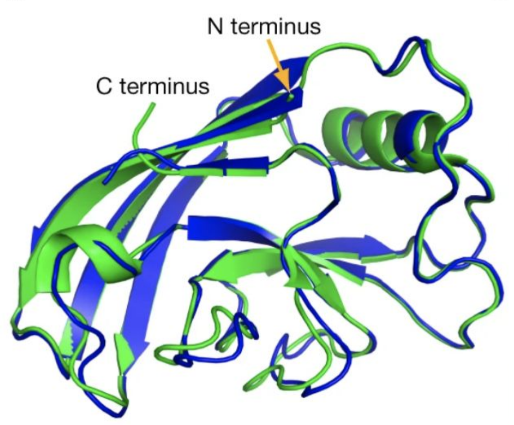


Image from: Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

Advances in deep learning 🤝 protein design 1/2

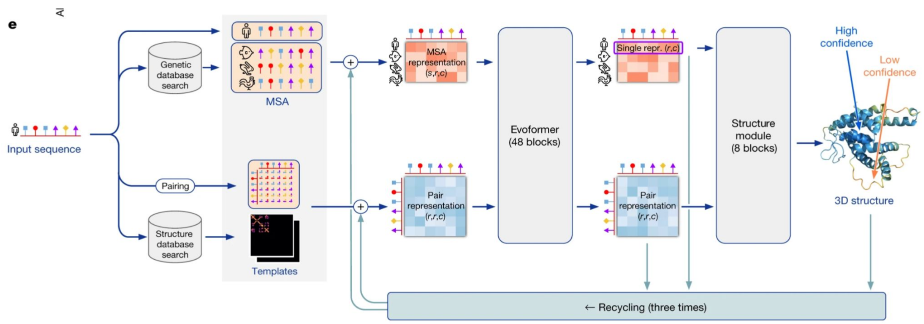
The protein folding revolution: AlphaFold2

- General solution to protein structure prediction
- Predicts accurate protein structures in minutes
  - As opposed to months - years for experimental determination



AlphaFold Experiment

r.m.s.d.$_{95}$ = 0.8 Å; TM-score = 0.93



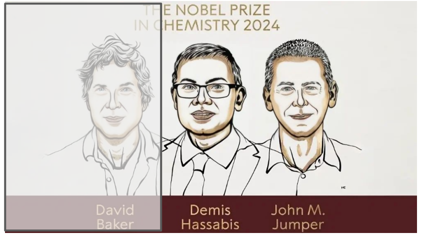
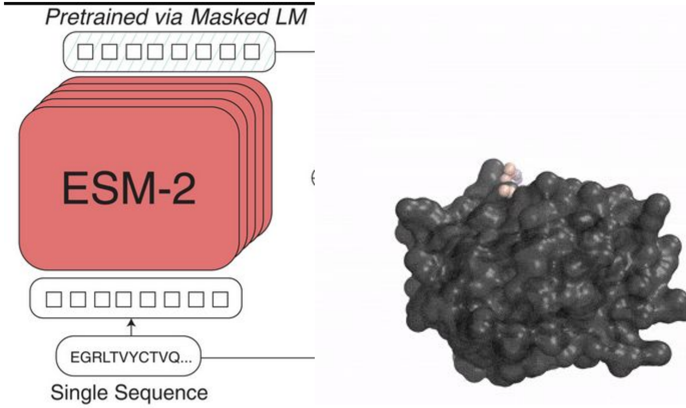
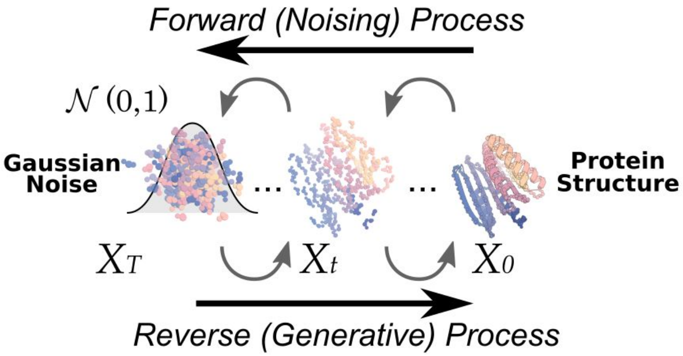


Image from: Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

Advances in deep learning 🤝 protein design 2/2

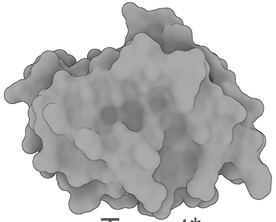
- DDPMs for denoising from a distribution
    - backbone design
- MPNNs for passing information along a graph
    - protein → graph, sequence → redesigned
- LLMs for capturing patterns in language
    - protein → language, evolution → pattern





Figures from: Ian Haydon, IPD @ UW; 10.1126/science.ade2574

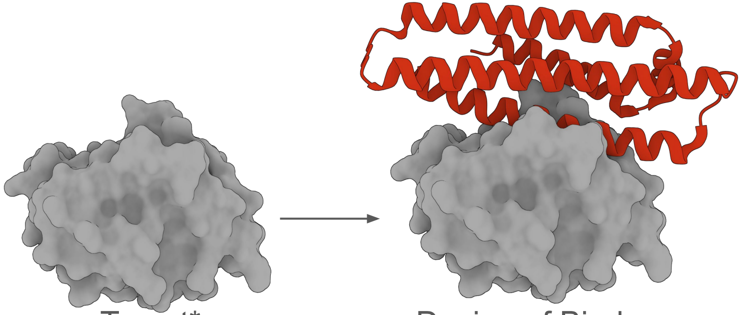
What is Protein Design - now



Target*

* Target can also be of small molecule or nucleic acid type

What is Protein Design - now



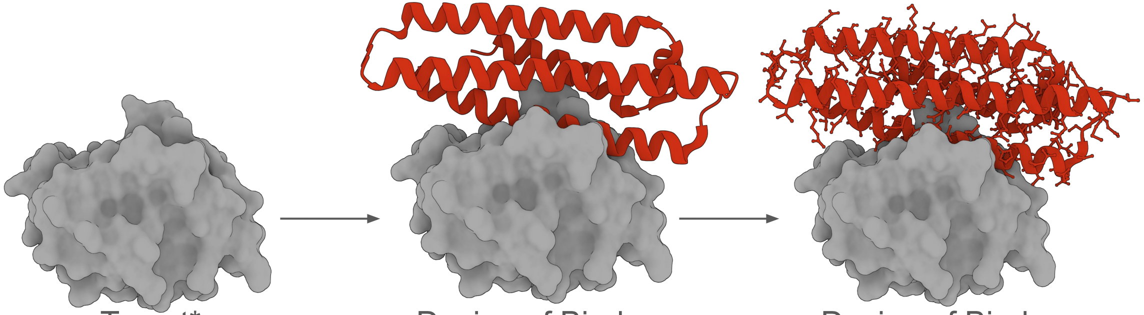
Target*

Design of Binder Backbone

RFDiffusion (allatom),
ESM,
AlphaProteo,
Chroma,
etc.

* Target can also be of small molecule or nucleic acid type

What is Protein Design - now



Target*

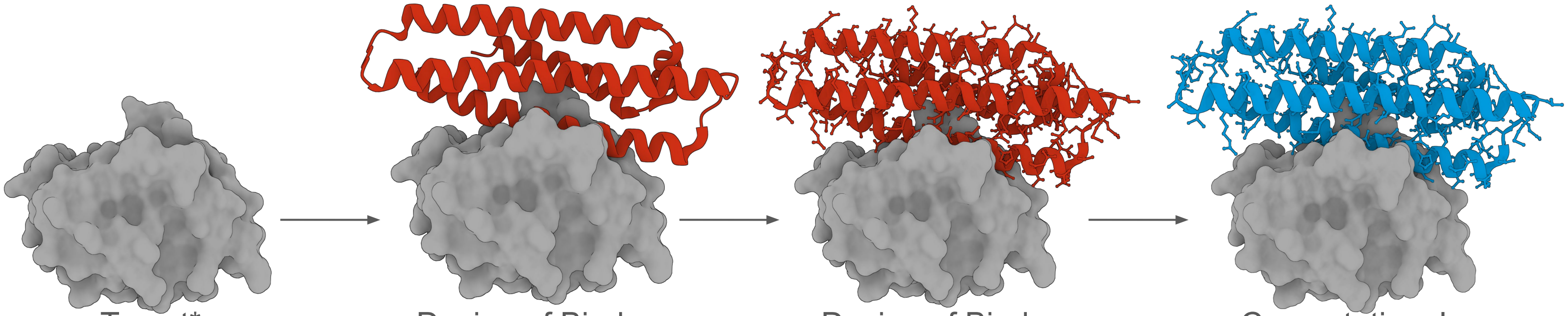
Design of Binder Backbone

Design of Binder Sequence

ProteinMPNN,
ESM-IF,
CARBonAra,
etc.

* Target can also be of small molecule or nucleic acid type

What is Protein Design - now



Target*

Design of Binder Backbone

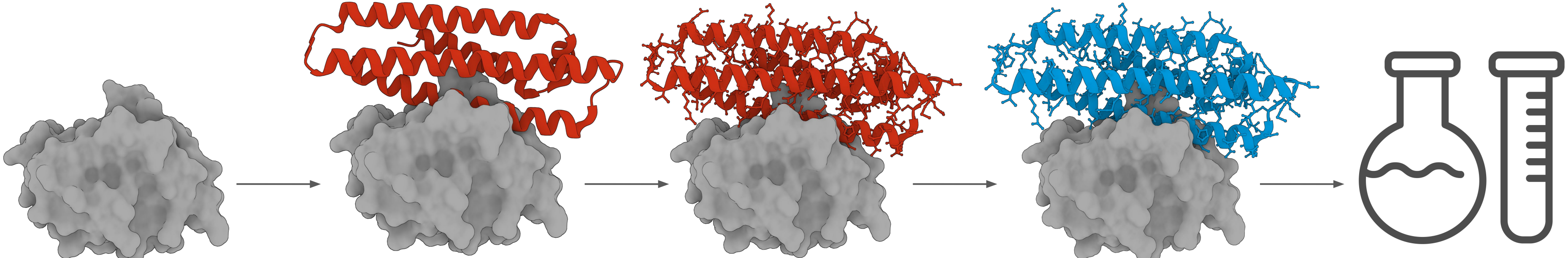
Design of Binder Sequence

Computational Evaluation

AlphaFold2/3,
RoseTTAfold2,
ESM2/3
Chai-1,
Boltz-1/2,
etc.

* Target can also be of small molecule or nucleic acid type

What is Protein Design - now



Target*

Design of Binder Backbone

Design of Binder Sequence

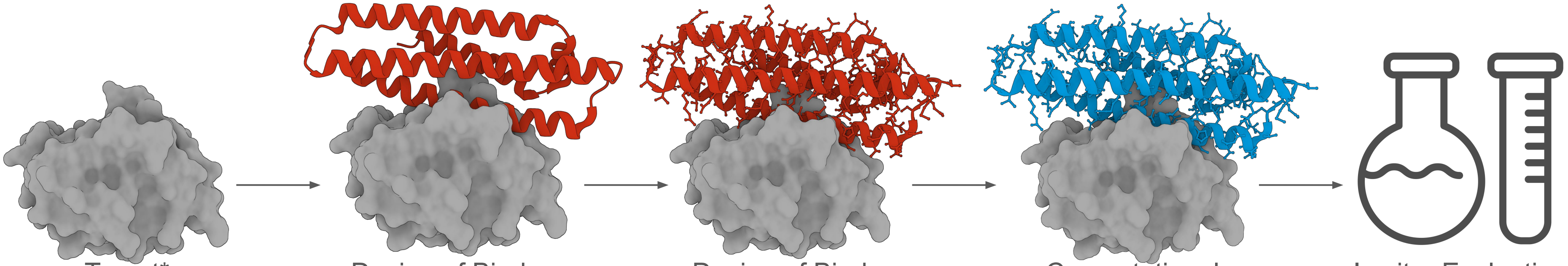
Computational Evaluation

In vitro Evaluation

CD,
ITC,
SPR,
Activity assay,
Cryo-EM,
XRD

* Target can also be of small molecule or nucleic acid type

What is Protein Design - now



Target*

Design of Binder Backbone

- RFDiffusion (allatom),
- ESM,
- AlphaProteo,
- Chroma,
- etc.

Design of Binder Sequence

- ProteinMPNN,
- ESM-IF,
- CARBonAra,
- etc.

Computational Evaluation

- AlphaFold2/3,
- RoseTTAfold2,
- ESM2/3
- Chai-1,
- Boltz-1/2,
- etc.

In vitro Evaluation

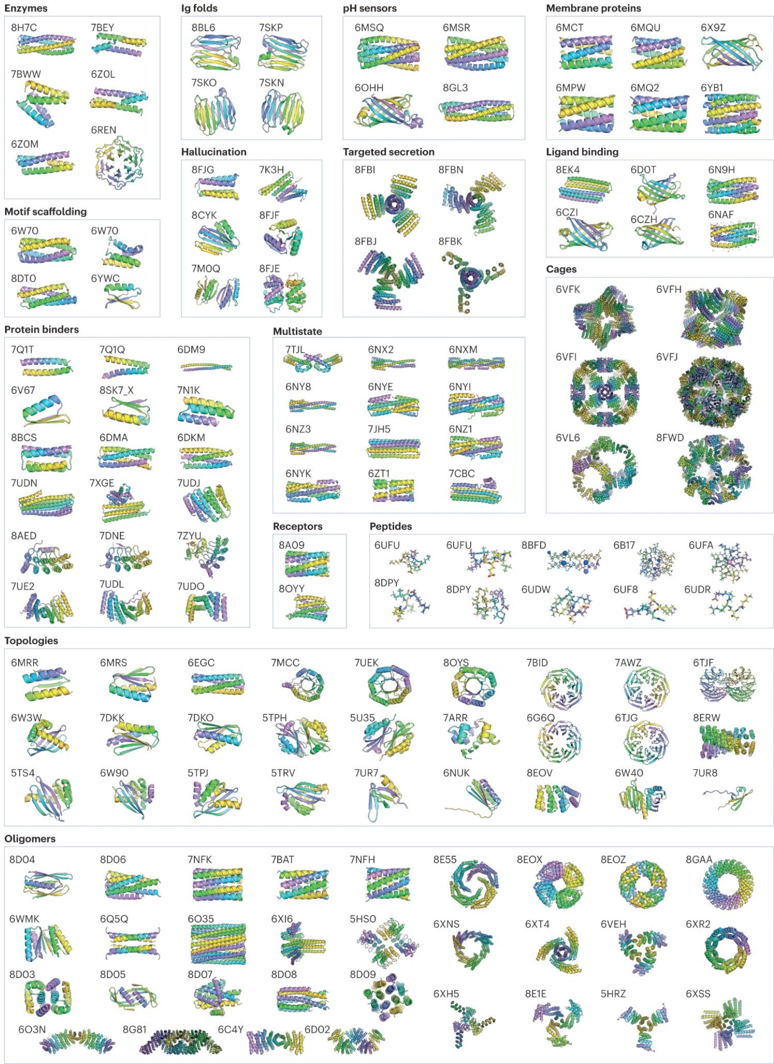
- CD,
- ITC,
- SPR,
- Activity assay,
- Cryo-EM,
- XRD

* Target can also be of small molecule or nucleic acid type

Showcase

Protein Design for everything

Awesome applications



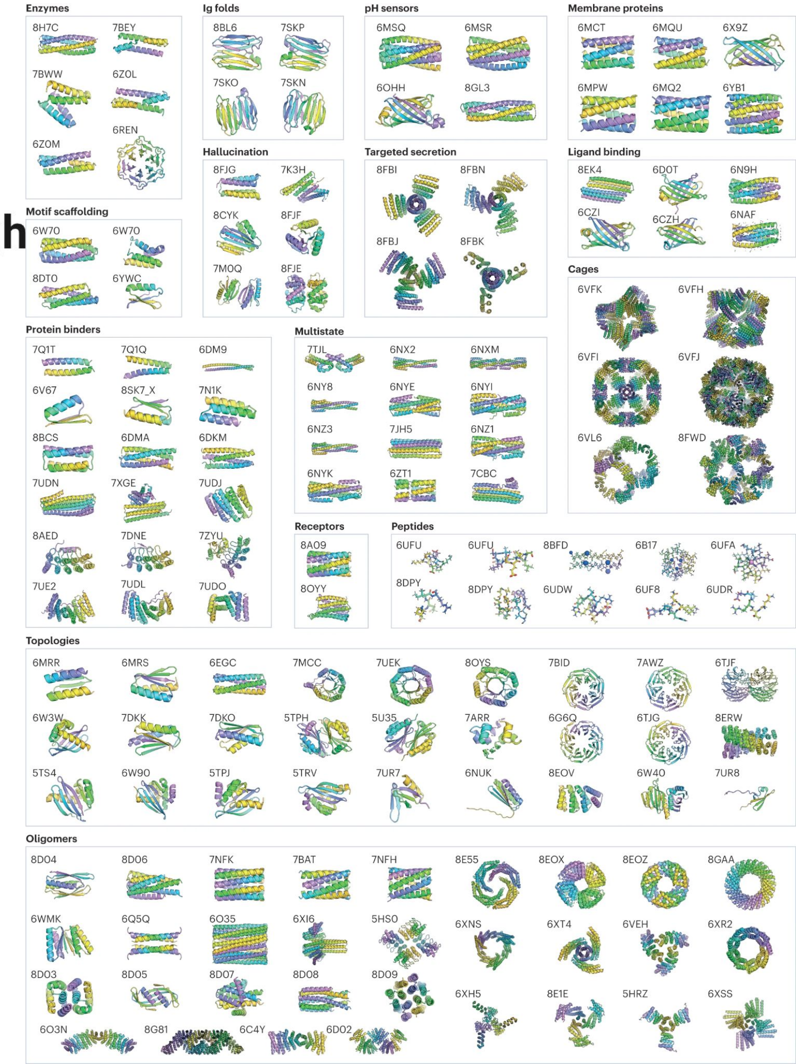
https://doi.org/10.1038/s41580-024-00718-y
https://www.nature.com/articles/s41587-024-02133-2/figures/6

Awesome applications

Sculpting conducting nanopore size and shape through de novo protein design

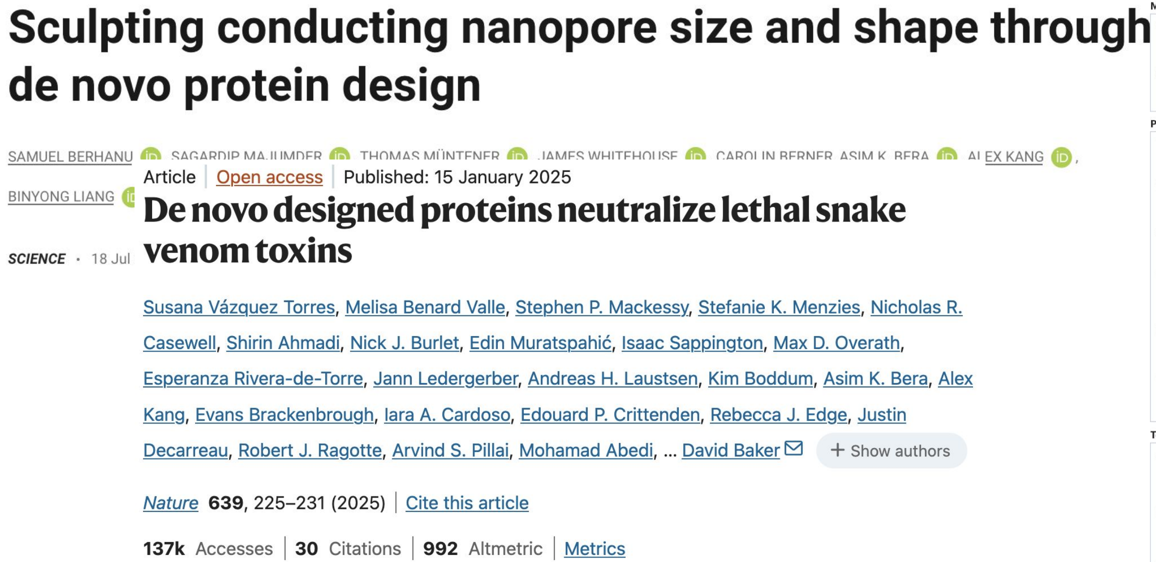
SAMUEL BERHANU, SAGARDIP MAJUMDER, THOMAS MÜNTENER, JAMES WHITEHOUSE, CAROLIN BERNER, ASIM K. BERA, ALEX KANG, BINYONG LIANG, NASIR KHAN, [...], AND ANASTASSIA A. VOROBIEVA +6 authors Authors Info & Affiliations

SCIENCE • 18 Jul 2024 • Vol 385, Issue 6706 • pp. 282-288 • DOI: 10.1126/science.adn3796



https://doi.org/10.1038/s41580-024-00718-y
https://www.nature.com/articles/s41587-024-02133-2/figures/6

Awesome applications



Sculpting conducting nanopore size and shape through de novo protein design

SAMUEL BERHANU, SAGARDIP MAJUMDER, THOMAS MÜNTENER, JAMES WHITEHOUSE, CAROLIN BERNER, ASIM K. BERA, ALEX KANG, BINYONG LIANG

SCIENCE • 18 Jul

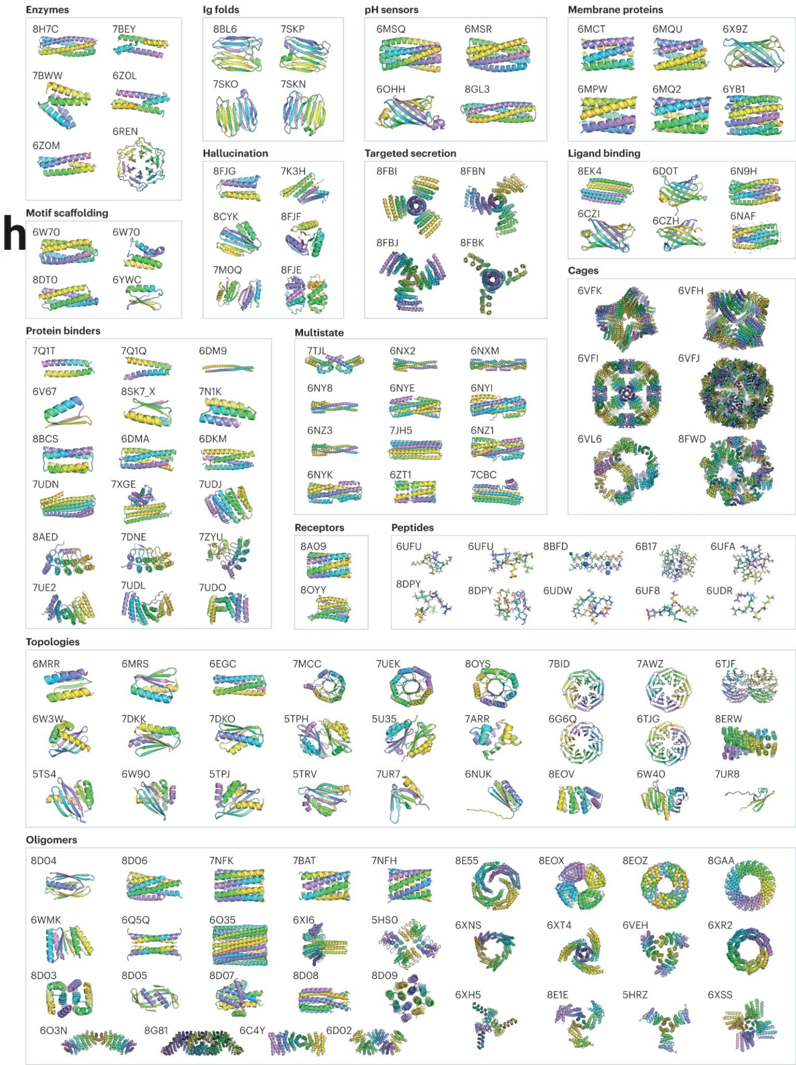
Article | Open access | Published: 15 January 2025

De novo designed proteins neutralize lethal snake venom toxins

Susana Vázquez Torres, Melisa Benard Valle, Stephen P. Mackessy, Stefanie K. Menzies, Nicholas R. Casewell, Shirin Ahmadi, Nick J. Burlet, Edin Muratspahić, Isaac Sappington, Max D. Overath, Esperanza Rivera-de-Torre, Jann Ledergerber, Andreas H. Laustsen, Kim Boddum, Asim K. Bera, Alex Kang, Evans Brackenbrough, Iara A. Cardoso, Edouard P. Crittenden, Rebecca J. Edge, Justin Decarreau, Robert J. Ragotte, Arvind S. Pillai, Mohamad Abedi, ... David Baker

Nature **639**, 225–231 (2025) | Cite this article

137k Accesses | **30** Citations | **992** Altmetric | Metrics



https://doi.org/10.1038/s41580-024-00718-y
https://www.nature.com/articles/s41587-024-02133-2/figures/6

Awesome applications

Sculpting conducting nanopore size and shape through de novo protein design

SAMUEL BERHANU, SAGARDIP MAJUMDER, THOMAS MÜNTENER, JAMES WHITEHOUSE, CAROLIN BERNER, ASIM K. BERA, ALEX KANG,

BINYONG LIANG

SCIENCE • 18 Jul

Article | Open access | Published: 15 January 2025

De novo designed proteins neutralize lethal snake venom toxins

Susana Vázquez Torres, Melisa Benar

Casewell, Shirin Ahmadi, Nick J. Burle

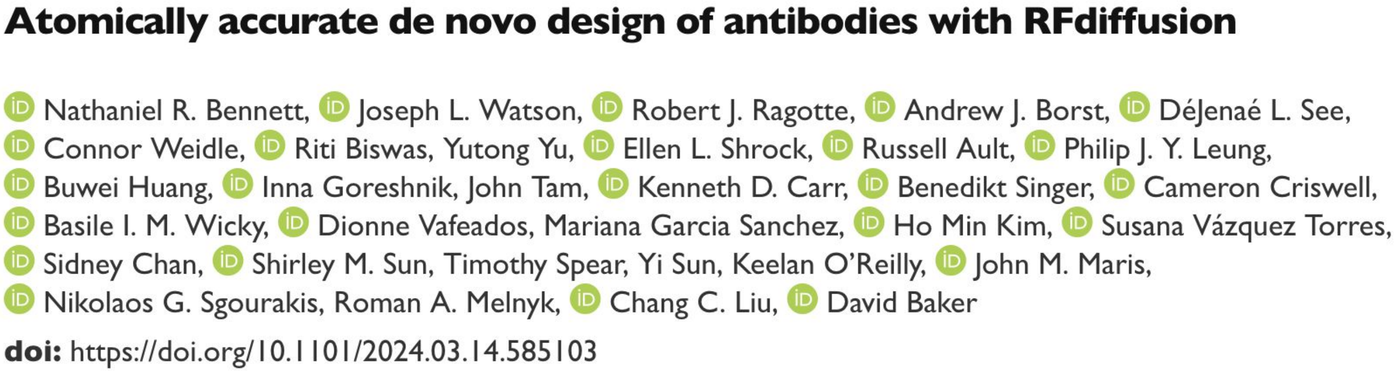
Esperanza Rivera-de-Torre, Jann Led

Kang, Evans Brackenbrough, Iara A. C

Decarreau, Robert J. Ragotte, Arvind

Nature **639**, 225–231 (2025) | Cite t

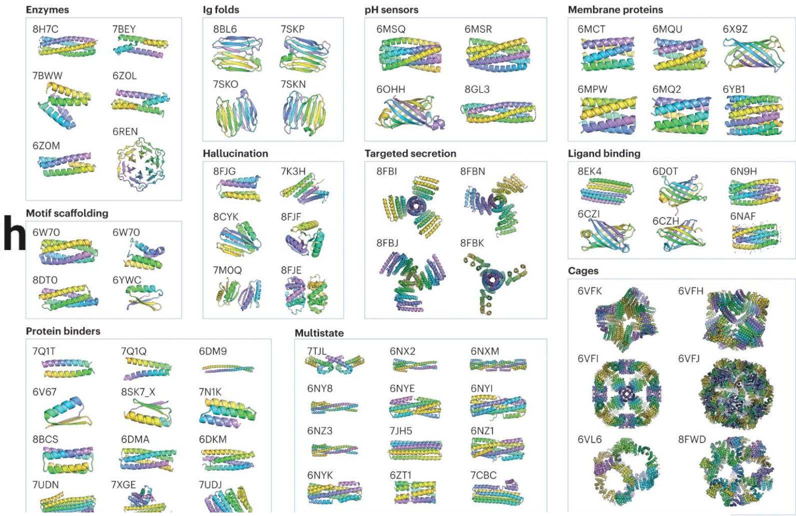
137k Accesses | **30** Citations | **99**

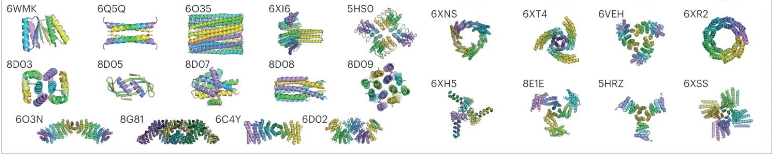


Atomically accurate de novo design of antibodies with RFdiffusion

Nathaniel R. Bennett, Joseph L. Watson, Robert J. Ragotte, Andrew J. Borst, DéJenaé L. See, Connor Weidle, Riti Biswas, Yutong Yu, Ellen L. Shrock, Russell Ault, Philip J. Y. Leung, Buwei Huang, Inna Goreshnik, John Tam, Kenneth D. Carr, Benedikt Singer, Cameron Criswell, Basile I. M. Wicky, Dionne Vafeados, Mariana Garcia Sanchez, Ho Min Kim, Susana Vázquez Torres, Sidney Chan, Shirley M. Sun, Timothy Spear, Yi Sun, Keelan O'Reilly, John M. Maris, Nikolaos G. Sgourakis, Roman A. Melnyk, Chang C. Liu, David Baker

doi: https://doi.org/10.1101/2024.03.14.585103





https://doi.org/10.1038/s41580-024-00718-y
https://www.nature.com/articles/s41587-024-02133-2/figures/6

Awesome applications

Sculpting conducting nanopore size and shape through de novo protein design

SAMUEL BERHANU, SAGARDIP MAJUMDER, THOMAS MÜNTENER, JAMES WHITEHOUSE, CAROLIN BERNER, ASIM K. BERA, ALEX KANG,

BINYONG LIANG

SCIENCE • 18 Jul

Article | Open access | Published: 15 January 2025

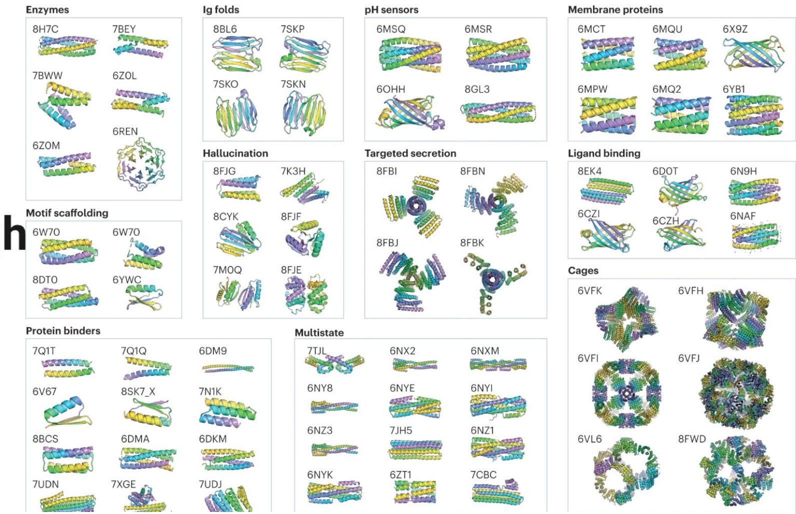
De novo designed proteins neutralize lethal snake venom toxins

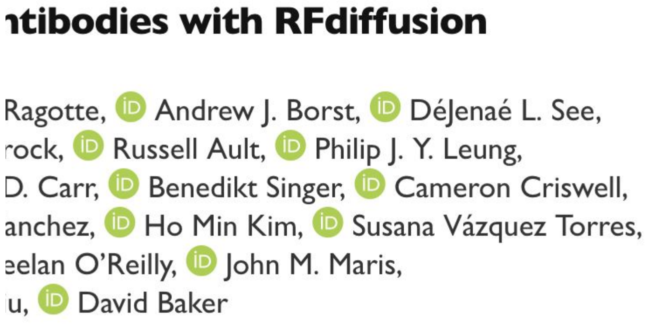
Susana Vázquez Torres, Melisa Benar

BindCraft: one-shot design of functional protein binders

Martin Pacesa, Lennart Nickel, Christian Schellhaas, Joseph Schmidt, Ekaterina Pyatova, Lucas Kissling, Patrick Barendse, Jagrity Choudhury, Srajan Kapoor, Ana Alcaraz-Serna, Yehlin Cho, Kourosh H. Ghamary, Laura Vinué, Brahm J. Yachnin, Andrew M. Wollacott, Stephen Buckley, Adrie H. Westphal, Simon Lindhoud, Sandrine Georgeon, Casper A. Goverde, Georgios N. Hatzopoulos, Pierre Gönczy, Yannick D. Muller, Gerald Schwank, Daan C. Swarts, Alex J. Vecchio, Bernard L. Schneider, Sergey Ovchinnikov, Bruno E. Correia

doi: https://doi.org/10.1101/2024.09.30.615802

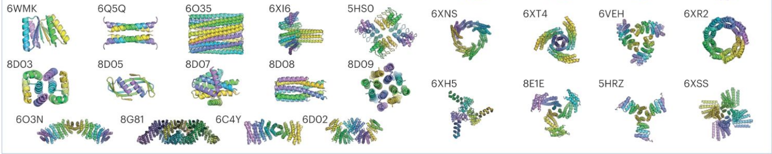




ntibodies with RFdiffusion

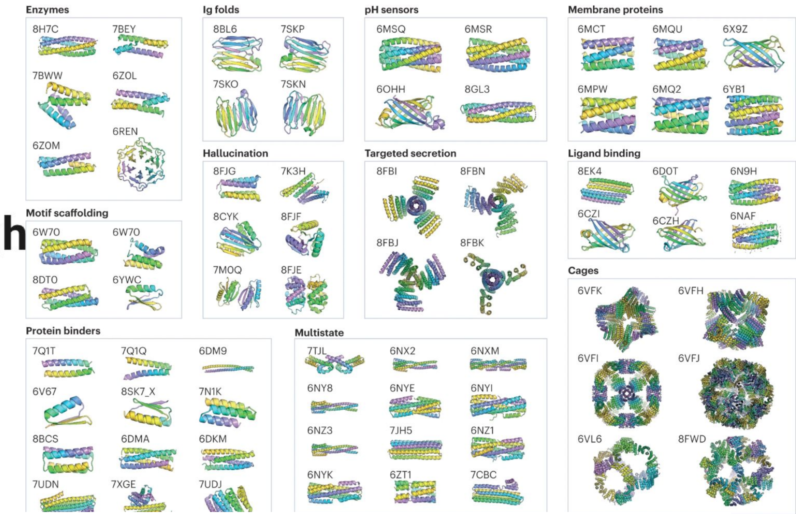
Ragotte, Andrew J. Borst, DéJenaé L. See,
rock, Russell Ault, Philip J. Y. Leung,
D. Carr, Benedikt Singer, Cameron Criswell,
anchez, Ho Min Kim, Susana Vázquez Torres,
eelan O'Reilly, John M. Maris,
u, David Baker

doi: https://doi.org/10.1101/2024.03.14.585103



https://doi.org/10.1038/s41580-024-00718-y
https://www.nature.com/articles/s41587-024-02133-2/figures/6

Awesome applications



Sculpting conducting nanopore size and shape through de novo protein design

SAMUEL BERHANU, SAGARDIP MAJUMDER, THOMAS MÜNTENER, JAMES WHITEHOUSE, CAROLIN BERNER, ASIM K. BERA, ALEX KANG,

BINYONG LIANG

SCIENCE · 18 Jul

Article | Open access | Published: 15 January 2025

De novo designed proteins neutralize lethal snake venom toxins

Susana Vázquez Torres, Melisa Benar

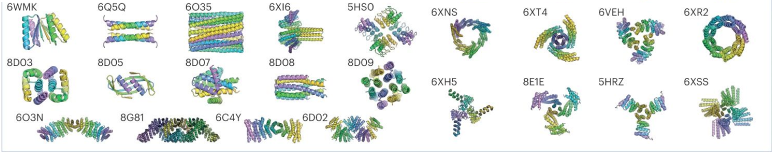
BindCraft: one-shot design of functional protein binders

ntibodies with RFdiffusion

Atom level enzyme active site scaffolding using RFdiffusion2

Woody Ahern, Jason Yim, Doug Tischer, Saman Salike, Seth M. Woodbury, Donghyo Kim, Indrek Kalvet, Yakov Kipnis, Brian Coventry, Han Raut Altae-Tran, Magnus Bauer, Regina Barzilay, Tommi S. Jaakkola, Rohith Krishna, David Baker

doi: https://doi.org/10.1101/2025.04.09.648075



https://doi.org/10.1038/s41580-024-00718-y
https://www.nature.com/articles/s41587-024-02133-2/figures/6

Your personal projects

Ideas and presentation of the available methods

What could a project look like

- Protein Design project
  - Apply a method/pipeline to design binders a specific targets
  - Analyse how different methods behave to different inputs for the same target
  - Example: Binder
- General computational structural biology project
  - Use methods on any structural question (i.e. docking, interfaces)
- Literature review

Different methods require different levels of expertise

Online

- easy-to-use (GUI, prompts)
- timeout when screen idles
- maybe not the most updated version
- resource limited (VRAM, space etc)

Izar

- only command line (maybe VSCode)
- no timeout (slurm)
- updates are easy (done by me)
- less resource limited

What is available online?

Whole pipelines:

- Colab Design (AFDesign + RFDiffusion)
- ProteinMPNN + Colabfold
- BindCraft
- Latent-Labs

Pure structure design:

- EvoDiff
- Chroma

What is available online?

Sequence redesign:

- Carbonara
- ProteinMPNN
- LigandMPNN

Structure prediction:

- AlphaFold3
- Colabfold
- Boltz webserver
- ESM3

What is available on izar?

- Pipelines
  - BindCraft
  - Boltzgen
- Structure generations
  - RFAntibody
  - RfDiffusion2
  - RfDiffusion
- Sequence redesign
  - Carbonara
  - ProteinMPNN
  - LigandMPNN

- Structure prediction
  - AlphaFold3
  - Chai-2
  - Boltz-2
- Other methods
  - conformere sampling w/ BioEmu
- What can I install if needed
  - Rosettafold3
  - RFpeptides
  - …

docker

CE

Possible ideas - protein design

PDB-IDs: 3ZTJ (H3), 3MJG (PDGFR), 4OGA (IR), 5U8R (IGF1R), 2GY7 (TIE2), 1XIW (CD3δ), 3KFD (TGFβ), 4O3V (VirB8) - Source: https://www.nature.com/articles/s41586-022-04654-9

KIT (PDB:2E9W), NRP1 (PDB:2QQI), PDGFRα (PDB:7LBF) - Source: rcsb.org ALK-2 (UniProt:Q04771), ALK-3 (UniProt:P36894), FCRL5 (UniProt:Q96RD9)

SARS-CoV-2 proteins (COVID-19 research): Spike glycoprotein: PDB IDs: 6VSB, 6VXX; Main protease: PDB ID: 5R7Y; NSP12 (RNA-dependent RNA polymerase): PDB ID: 6XEZ; NSP16 (2'-O-methyltransferase): PDB ID: 6W4H; ORF8: PDB ID: 7JTL

Cancer-related proteins: Bromodomain-containing protein 4 (BRD4): UniProt ID: O60885, PDB ID: 8RX0; Von Hippel-Lindau disease tumor suppressor (VHL): UniProt ID: P40337, PDB ID: 8RX0;

Ubiquitin-proteasome system:Ubiquitin-conjugating enzyme E2 R1 (UBE2R1): UniProt ID: P49427, PDB ID: 8RX0; E3 ubiquitin-protein ligase RBX1: UniProt ID: P62877, PDB ID: 8RX0

Source: https://www.biorxiv.org/content/10.1101/2024.10.11.617496v2.full

Small review

- Pick a protein system, pathway, interaction network etc.
- Summarize the computational structural studies
- Describe the methods and how they integrate